View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by OpenEdition

Mathématiques et sciences humaines
Mathematics and social sciences

199 | 2012
Psychologie et mathématiques

Rating the diversity in sets of objects by referring to transformations as criteria

Évaluer la diversité dans des ensembles d'objets en faisant référence aux transformations comme critères

Luigi Burigana and Michele Vicovaro

OpenEdition Journals

Electronic version
URL: http://journals.openedition.org/msh/12268
DOI: 10.4000/msh.12268
ISSN: 1950-6821

Publisher
Centre d'analyse et de mathématique sociales de l'EHESS

Printed version
Date of publication: 15 September 2012
Number of pages: 27-44
ISSN: 0987-6936

Electronic reference
Luigi Burigana and Michele Vicovaro, « Rating the diversity in sets of objects by referring to transformations as criteria », *Mathématiques et sciences humaines* [Online], 199 | 2012, Online since 04 December 2012, connection on 01 May 2019. URL : http://journals.openedition.org/msh/12268 ; DOI : 10.4000/msh.12268

© École des hautes études en sciences sociales

Math. Sci. hum. / Mathematics and Social Sciences (50e année, n° 199, *2012*(3), p. 27–44)

RATING THE DIVERSITY IN SETS OF OBJECTS BY REFERRING TO TRANSFORMATIONS AS CRITERIA

Luigi BURIGANA[1], Michele VICOVARO[1]

RÉSUMÉ – Évaluer la diversité dans des ensembles d'objets en faisant référence aux transformations comme critères
Le point de départ de cette étude est la définition de deux concepts relatant comment un ensemble de transformations, agissant à l'intérieur d'un domaine, peut représenter une limite supérieure ou inférieure à la diversité existante dans n'importe quel sous-ensemble de ce domaine. Le sujet de cette analyse est ensuite graduellement étendu, considérant des partitions (plutôt que des sous-ensembles) d'un domaine, des familles d'ensembles de transformations (plutôt qu'un seul ensemble), et du genre des ensembles d'objets indirectement reliés à ce domaine (plutôt que des ensembles directement inclus dans le domaine). Les principaux concepts définis sont explorés dans leurs propriétés formelles et illustrés au travers d'exemples. Les sections introductive et conclusive incluent des commentaires sur la motivation et les avantages possibles de la méthode discutée.

MOTS-CLÉS – Diversité, Invariance, Transformation interne

SUMMARY – *The starting point of this study is the definition of two concepts expressing how a set of transformations acting within a domain may represent an upper or a lower bound of the diversity existing in any subset of that domain. The subject of analysis is then gradually expanded, by considering partitions (rather than subsets) of the domain, families of sets of transformations (rather than just one such set), and sets of objects indirectly related to the domain (rather than sets directly included in it). The main concepts defined in this study are explored in their formal properties and illustrated by examples. The introductory and concluding sections include comments on the motivation for the study and the possible merits of the method discussed.*

KEYWORDS – Diversity, Inner transformation, Invariance

1. INTRODUCTION

Similarity and its antonym dissimilarity are supposed to play important roles in various cognitive processes, such as categorization, learning, visual search, social judgement, etc. In consequence, the construction of methods for evaluating those relationships forms an important topic of cognitive psychology. A distinction may be made between *empirical methods*, which estimate similarity and dissimilarity based on the answers of real subjects in suitable experiments, and *analytic methods*, which evaluate those relationships by analysing the properties implicit in the objects to be compared, or assessing their position relative to an *a priori* criterion. The classic "contrast model" of Tversky [1977] and the recent use of the "representational distortion" criterion of Hahn, Chater, Richardson [2003] are examples of the analytic method. In particular, this criterion is described

[1]Dipartimento di Psicologia Generale, Università di Padova, Via Venezia 8, I-35131 Padova (Italy), luigi.burigana@unipd.it, vicovaro85@gmail.com

as a "transformational approach to similarity", because of the part played in it by the concept of transformation. Specifically, a set T of "basic transformations" is presumed, and these transformations are understood as production rules (i.e., ways of constructing new objects from given objects by additions, deletions, changes, etc.); the dissimilarity of an object y from an object x is computed as the minimum length of a sequence of transformations in T which is able to produce y starting from x (the so-called "transformational distance" of y from x).

The method we discuss in this paper is itself "analytic" and "transformational" in the generic senses stated above. However, it differs markedly from "representational distortion" in two respects. The first concerns the target of the method: in an application, a set A of objects is considered, and the aim is to rate the *overall diversity* within A (rather than the dissimilarity between its objects taken in pairs). In the concluding section we mention possible research situations in which such a target may have scientific importance. Because the word "dissimilarity" is regularly used in psychometrics for pairwise measures of lack of similarity (e.g. [Barthélemy, Guénoche, 1991]) we use the word "diversity" here for lack of similarity as the collective property of a set of objects. The second respect concerns the transformations in a reference set T (or a hierarchy $T_1 \subset \ldots \subset T_m$ of such sets), which in our method intervene as *substitution rules* (rather than production rules). More precisely, we hypothesize a basic set X of objects (of which the set A to be measured may be a subset), and a set T of transformations acting within X (i.e., T is a subset of X^X, the set of all functions having X as domain and codomain). Each transformation $t \in T$ is interpreted here as a substitution rule, so that the equation $y = t(x)$ (for any $x, y \in X$) means that object y is similar or adjacent to object x (according to the criterion represented by t). More generally, for any $x, y \in X$, the existence of a $t \in T$ such that $y = t(x)$ means that y is similar or adjacent to x (according to the criterion collectively represented by the set of transformations T). We must emphasize that the transformations referred to in this study are defined over the *whole* basic set X, and that this assumption is essential for a special property of our method, that is, the property that the same transformation set T may serve as a criterion for judging similarity and diversity not only within sets of elements of X, but also within sets of objects of higher complexity which are constructed using elements of X. For example, a triple (y_1, y_2, y_3) of elements of X may be judged to be similar or adjacent to another triple (x_1, x_2, x_3) of elements of X if $y_1 = t(x_1)$ and $y_2 = t(x_2)$ and $y_3 = t(x_3)$ for some $t \in T$ (the *same* t for all three components). This issue is developed in Section 5.

The idea of referring to a set T of transformations within a domain X as a set of legitimate substitution rules replicates a standard notion associated with permutation groups – for example, the group of automorphisms of a structure is the set of substitution rules on its domain which leave the structure unchanged. Moreover, the idea of referring to a hierarchy $T_1 \subset \ldots \subset T_m$ of such transformation sets as a reference system for rating the diversity within sets of objects is itself analogous to a known kind of application of permutation groups (our reference here is to theories imitating the so-called "Erlanger Program" of theoretical geometry, to be mentioned in Section 4). Actually, owing to these analogies, there are points in this paper in which we refer to elementary concepts of the theory of permutation groups, in particular the concepts of "invariant set", "orbit", and "action" of a group. For ease of reference, let us here recall the meanings of these concepts. If T is a permutation group on a domain X (that is, a set of bijective functions from X onto X, closed under composition and inversion) and $A \subseteq X$, then A is *invariant* under T if $t(A) = A$ for all $t \in T$ (where $t(A) = \{t(x) : x \in A\}$). A set $A \subseteq X$ is a T-*orbit* if $T(x) = A$ for all $x \in A$ (where $T(x) = \{t(x) : t \in T\}$). The *action* of the group relates to the fact that each permutation $t \in T$, while directly acting within the domain X, may indirectly also act on other domains if these are systematically related to X (for example, any $t \in T$ induces a permutation t^* on the power set 2^X by the rule $t^*(\{x_1, \ldots, x_n\}) = \{t(x_1), \ldots, t(x_n)\}$ for all $\{x_1, \ldots, x_n\} \in 2^X$, so that t acts on 2^X as made explicit by t^*) [Dixon, Mortimer, 1996, pp. 5-6].

Owing to the analogies stated above, the contexts for the possible use of the method in this study are similar to those for the use of permutation groups for classification purposes involving the invariance condition. There are, however, two noticeable features distinguishing our approach in this regard. One is the definition of a reference transformation set T, which we presume to be *any set* of inner transformations of a basic set X, not necessarily a permutation group (i.e., not necessarily a set closed to inversion and composition, nor even a set of bijective transformations). The main reason for this choice of greater freedom is that, in our approach, set T determines a similarity on X, and when T is a permutation group then the similarity it determines is necessarily an equivalence (the quotient set of X modulo that equivalence is the partition of X into T-orbits). But it is well known in cognitive psychology that similarity relationships as revealed by psychological experiments may fail to be equivalences, that is, may fail to be symmetric and transitive relations [Tversky, 1977, pp. 328-329][2]. Thus in order to model psychological similarities and dissimilarities accurately, we must suspend the general assumption that T is a permutation group, and only presume that it is a set of inner transformations of the basic set X. Our formal analysis in this paper amounts to an exploration of the regularities which still hold true in such less constrained conditions. The other special feature of our approach is the distinction between *two ways* in which an object set A may relate to a transformation set T, which we denote by $T \rhd A$ and $T \lhd A$, and describe by stating that T represents, respectively, an "upper bound" and a "lower bound" of the diversity in A. These two concepts are formally defined and examined in their basic properties in Section 2. This second feature (the distinction between the relations $\rhd$ and $\lhd$) is a consequence of the first one (the reference to generic sets of inner transformations) using the following argument. If T is a permutation group, then it is seen that the sets A such that $T \rhd A$ are the subsets of T-orbits, and the sets A such that $T \lhd A$ are the subsets of X expressible as unions of T-orbits. Thus, when T is a permutation group, there is a simple direct connection between the two kinds of sets, linked by the orbits of the group. Otherwise, when T is a generic set of inner transformations (not a permutation group), then the relationship between the two kinds of sets becomes looser (since there is no family of orbits serving as a common basis), and each of them deserves its own analysis which constitutes our main task in the following sections.

The plan of the paper is as follows. In Section 2 we discuss the basic case, when one object set $A \subseteq X$ becomes related to one transformation set $T \subseteq X^X$. In Section 3 we make a first step towards complexity, by presuming that the object side of the paradigm is a family $\mathcal{P} = \{P_1, \ldots, P_n\}$ of subsets of X (specifically, a partition of X). In Section 4 we make a further step towards complexity, by presuming that the transformation side is an ordered family $\mathcal{T} = (T_1, \ldots, T_m)$ of transformation sets. In Section 5 we discuss a form of "indirect action" of transformation sets, so that transformations acting directly within a domain X may serve as criteria for judging diversity within another domain X^*. In Section 6 we comment on the possible uses and merits of the transformational approach we have described.

2. A SET OF TRANSFORMATIONS AND A SUBSET OF THEIR DOMAIN

Throughout our discussion we assume that the basic set X has *finite* cardinality.

DEFINITION 1. *Let X, $T \subseteq X^X$, and $A \subseteq X$ be as presumed so far. The transformation set T represents an* upper bound *(symbol $T \rhd A$) or* lower bound *(symbol $T \lhd A$) of the diversity within the object set A depending on whether the first or the second of the following two conditions holds true:*

$$T(x) \supseteq A \text{ for all } x \in A$$
$$T(x) \subseteq A \text{ for all } x \in A$$

in which $T(x) = \{t(x) : t \in T\}$.

The meaning of these concepts may be illustrated by considering the following alternative expressions:

$$T \rhd A \text{ iff } (\forall x \in A)(\forall y \in A)(\exists t \in T)(y = t(x))$$
$$T \lhd A \text{ iff } (\forall x \in A)(\forall t \in T)(\exists y \in A)(y = t(x)).$$

Thus, $T \rhd A$ means that for any two objects x and y in A there is some transformation t in T which carries from x to y, so that the overall diversity within A is covered by the diversity expressible

[2]There are also standard sets of transformations endowed with scientific importance which fail to be groups or semigroups. The set of "perspective transformations" referred to in perceptual psychology and computer vision is an example [Mundy, Zisserman, 1992, p. 475].

through T. Conversely, $T \lhd A$ means that for any object x in A and transformation t in T, by applying t to x an object $t(x)$ is obtained which is itself in A, so that the diversity expressible through T is covered by the overall diversity within A. As a special case, if T is a permutation group (on a finite domain X), then $T \lhd A$ means that $t(A) = A$ for all $t \in T$ (i.e., A is invariant under T), and the conjunction ($T \rhd A$ and $T \lhd A$) means that $T(x) = A$ for all $x \in A$ (i.e., A is a T-orbit).

A convenient way of expressing the concepts in Definition 1 is by means of a binary relation on X defined in this way:

$$R = \{(x, y) \in X^2 : y = t(x) \text{ for some } t \in T\}.$$

We call R the *directed adjacency* determined by transformation set T. Henceforth we presume that R is a reflexive relation, which is certainly true if the identity transformation id_X on the domain X belongs to T. Moreover, for any $A \subseteq X$, by $R(A)$ we denote the image of set A through the relation R, so that:

$$R(A) = \{y \in X : y = t(x) \text{ for some } x \in A \text{ and } t \in T\}.$$

PROPOSITION 1. *Let X, $T \subseteq X^X$, $R \subseteq X^2$, $A \subseteq X$, and $R(A) \subseteq X$ be as stated above.*
(i) $T \rhd A$ if and only if $R \supseteq A^2$.
(ii) $T \lhd A$ if and only if $R(A) \subseteq A$.

Proof. (i) If $T \rhd A$ and $(x, y) \in A^2$, then $y \in A \subseteq T(x)$, so that $y = t(x)$ for some $t \in T$, and $(x, y) \in R$. If $R \supseteq A^2$ and $x, y \in A$, then $(x, y) \in R$, so that $y = t(x)$ for some $t \in T$, and $y \in T(x)$. (ii) If $T \lhd A$ and $y \in R(A)$, then $y = t(x) \in T(x) \subseteq A$ for some $x \in A$ and $t \in T$, so that $y \in A$. If $R(A) \subseteq A$, $x \in A$ and $y \in T(x)$, then $y = t(x)$ for some $t \in T$, so that $y \in R(A)$, and $y \in A$. $\square$

As an example, let us consider $X = \{a, b, \ldots, l\}$ as the basic set, and a set $T = \{t_1, \ldots, t_6\}$ of 6 inner transformations defined as follows (where t_1 is the identity id_X):

$$\begin{aligned}
t_1 &= \{(a,a),(b,b),(c,c),(d,d),(e,e),(f,f),(g,g),(h,h),(i,i),(j,j),(k,k),(l,l)\} \\
t_2 &= \{(a,b),(b,c),(c,d),(d,b),(e,f),(f,f),(g,h),(h,i),(i,j),(j,h),(k,k),(l,k)\} \\
t_3 &= \{(a,a),(b,d),(c,c),(d,c),(e,e),(f,e),(g,i),(h,j),(i,g),(j,j),(k,l),(l,l)\} \\
t_4 &= \{(a,b),(b,b),(c,b),(d,d),(e,d),(f,e),(g,g),(h,g),(i,h),(j,h),(k,j),(l,h)\} \\
t_5 &= \{(a,a),(b,c),(c,b),(d,b),(e,b),(f,f),(g,i),(h,i),(i,h),(j,i),(k,l),(l,k)\} \\
t_6 &= \{(a,b),(b,d),(c,c),(d,c),(e,f),(f,e),(g,h),(h,j),(i,g),(j,i),(k,k),(l,h)\}.
\end{aligned} \tag{1}$$

The directed adjacency determined by T is described by the following list of pairs and is represented in Figure 1:

$$\begin{aligned}
R = \{&(a,a),\ldots,(l,l),(a,b),(b,c),(b,d),(c,b),(c,d),(d,b),(d,c),(e,b),(e,d),(e,f),(f,e), \\
&(g,h),(g,i),(h,g),(h,i),(h,j),(i,g),(i,h),(i,j),(j,h),(j,i),(k,j),(k,l),(l,h),(l,k)\}.
\end{aligned}$$

Then consider these subsets of X : $A_1 = \{b, c, d\}$, $A_2 = \{g, h, i\}$, $A_3 = \{a, b, c, d\}$, and $A_4 = \{a, c, e, g\}$. By applying Proposition 1 we see that $T \rhd A_1$ and $T \lhd A_1$; $T \rhd A_2$ but not $T \lhd A_2$; $T \lhd A_3$ but not $T \rhd A_3$; and neither $T \rhd A_4$ nor $T \lhd A_4$.

Owing to Proposition 1, the relations $\rhd$ and $\lhd$ depend on a transformation set through the directed adjacency it determines. For this reason, and for the immediacy of our proofs, in stating the basic properties of both relations we find it preferable to refer to the adjacency R (rather than the transformation set T determining that adjacency)[3]. A list of elementary properties is given in the next proposition. For stating one of them we apply the following special concept: a relation $R \subseteq X^2$ is *part-wise symmetric* if for all $A \subseteq X$, if there are $x \in A$ and $y \in X \setminus A$ such that

[3]However, there are aspects of the action of T that cannot be mediated by R, such as any process of "transformation induction" as discussed in Section 5. Thus, the initial concept T (the set of inner transformations) cannot be removed from our analysis.

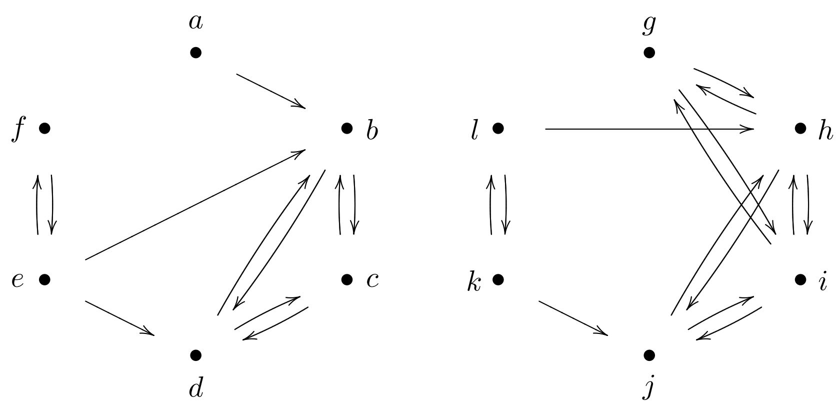

FIGURE 1. The directed adjacency determined by a set of 6 transformations on a set of 12 points (loops at the points are implied).

$(x, y) \in R$, then there are also $x' \in A$ and $y' \in X \setminus A$ such that $(y', x') \in R$. This amounts to a weakening of the standard symmetry property of binary relations[4].

PROPOSITION 2. *For all $T \subseteq X^X$, the resulting adjacency $R \subseteq X^2$, and all $A, B \subseteq X$ these implications hold true:*
(i) if $T \rhd A$ and $A \supseteq B$, then $T \rhd B$;
(ii) if R is transitive, $A \cap B \neq \emptyset$, $T \rhd A$, and $T \rhd B$, then $T \rhd A \cup B$;
(iii) if $T \lhd A$ and $T \lhd B$, then $T \lhd A \cap B$ and $T \lhd A \cup B$;
(iv) if R is part-wise symmetric and $T \lhd A$, then $T \lhd X \setminus A$;
(v) if $T \lhd A$, then $T \lhd R(A)$;
(vi) if $T \rhd A$ and $T \lhd B$, then either $A \subseteq B$ or $A \cap B = \emptyset$.

Proof. (i) This is implied by Proposition 1.i. (ii) Presume the four hypotheses and consider any $x, y \in A \cup B$. If $x, y \in A$ or $x, y \in B$, then $(x, y) \in A^2 \cup B^2 \subseteq R$. If $x \in A \setminus B$ and $y \in B \setminus A$, then $(x, z), (z, y) \in A^2 \cup B^2 \subseteq R$ for any $z \in A \cap B \neq \emptyset$, so that $(x, y) \in R$. Thus $(A \cup B)^2 \subseteq R$, and $T \rhd A \cup B$ by Proposition 1.i. (iii) If $T \lhd A$ and $T \lhd B$, then $R(A) \subseteq A$ and $R(B) \subseteq B$ by Proposition 1.ii, hence $R(A \cap B) \subseteq R(A) \cap R(B) \subseteq A \cap B$ and $R(A \cup B) = R(A) \cup R(B) \subseteq A \cup B$, so that $T \lhd A \cap B$ and $T \lhd A \cup B$ by the same proposition. (iv) If $T \lhd A$, then $R(A) \subseteq A$, so that there are no $x \in A$ and $y \in X \setminus A$ such that $(x, y) \in R$. Part-wise symmetry of R implies that neither can there be any $x' \in A$ and $y' \in X \setminus A$ such that $(y', x') \in R$, thus $R(X \setminus A) \subseteq X \setminus A$, and $T \lhd X \setminus A$. (v) If $T \lhd A$, then $R(A) \subseteq A$, hence $R(R(A)) \subseteq R(A)$, and $T \lhd R(A)$. (vi) Suppose (ad absurdum) $T \rhd A$, $T \lhd B$, $A \cap B \neq \emptyset$, but not $A \subseteq B$, and take any $x \in A \cap B$ and $y \in A \setminus B$. Thus $(x, y) \in A^2 \subseteq R$, $x \in B$, $y \notin B$, and $y \in R(B) \setminus B$, which contradicts the hypothesis $T \lhd B$. $\square$

The next proposition highlights some simple properties of the relations $\rhd$ and $\lhd$ which depend on the choice of the transformation set. In the fourth property reference is made to a *semigroup* as a set of inner transformations which is closed under composition (symbol $\circ$).

PROPOSITION 3. *For all $T, S \subseteq X^X$, the adjacencies $R_T, R_S \subseteq X^2$ resulting from them, and all $A \subseteq X$ these implications are true:*
(i) if $T \rhd A$ and $R_T \subseteq R_S$, then $S \rhd A$;
(ii) if $T \lhd A$ and $R_T \supseteq R_S$, then $S \lhd A$;
(iii) if $T \lhd A$ and $S \lhd A$, then $T \cup S \lhd A$;
(iv) if T is the semigroup generated by S, then $T \lhd A$ if and only if $S \lhd A$.

[4]This characterization holds true: a binary relation R is part-wise symmetric on domain X if and only if the strong components of the digraph (X, R) are completely isolated from one another (there is no line between distinct strong components). We thank one anonymous reviewer for pointing out this property.

Proof. Parts (i) and (ii) are implied by Proposition 1. (iii) If $T \lhd A$ and $S \lhd A$, then $R_T(A) \subseteq A$ and $R_S(A) \subseteq A$, so that $R_{T\cup S}(A) = (R_T \cup R_S)(A) = R_T(A) \cup R_S(A) \subseteq A \cup A = A$, and $T \cup S \lhd A$ by Proposition 1.ii. (iv) The "only if" part follows from part (ii) of this proposition, as $T \supseteq S$. For proving the "if" part, suppose $S \lhd A$ and consider any $x \in A$ and $t \in T$, so that $t = s_k \circ \cdots \circ s_1$ for some $s_1, \ldots, s_k \in S$ (as T is generated by S). Then $t(x) = s_k(\ldots s_1(x) \ldots) \in A$ (because $S \lhd A$). This is true of all $x \in A$ and $t \in T$, so that $T \lhd A$ by Definition 1. □

For any fixed transformation set $T \subseteq X^X$, let ${}^{\rhd}\mathcal{A}$ and ${}^{\lhd}\mathcal{A}$ be the families of subsets of X that have T as an upper and, respectively, lower bound of diversity:

$$
\begin{aligned}
{}^{\rhd}\mathcal{A} &= \{A \in 2^X : T \rhd A\} \\
{}^{\lhd}\mathcal{A} &= \{A \in 2^X : T \lhd A\}.
\end{aligned}
$$

Figure 2 represents (as a unified Hasse diagram) the families ${}^{\rhd}\mathcal{A}$ and ${}^{\lhd}\mathcal{A}$ resulting from T described by (1).

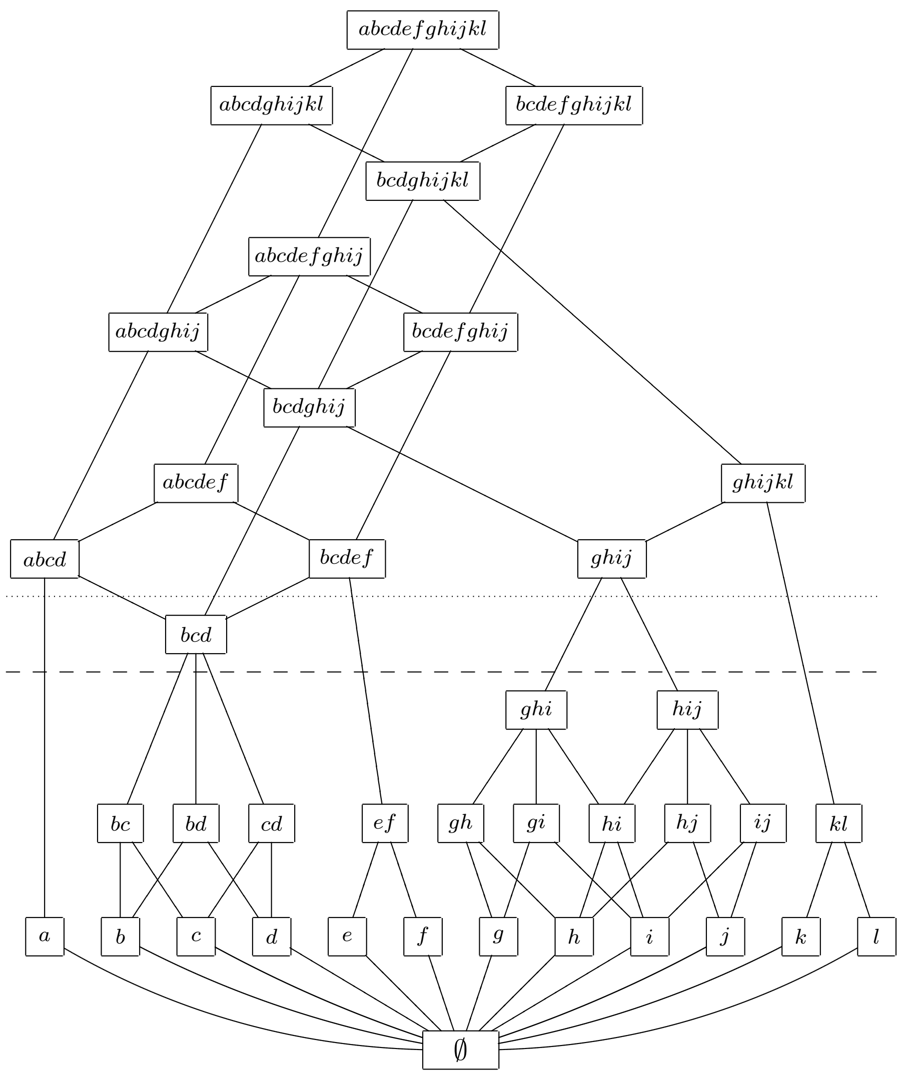

FIGURE 2. Hasse diagram of the families ${}^{\rhd}\mathcal{A}$ and ${}^{\lhd}\mathcal{A}$ resulting from the transformation set described by (1). Members of ${}^{\rhd}\mathcal{A}$ are the sets below the dotted line, and members of ${}^{\lhd}\mathcal{A}$ are the sets above the dashed line (including $\emptyset$).

The next proposition specifies the algebraic profiles of both families. In the proposition reference is made to the directed adjacency R determined by T, the *symmetric part* of R ($\mathrm{sp}(R) = R \cap R^{-1}$, where R^{-1} is the reverse of R), and the *transitive closure* of R ($\mathrm{tc}(R) = R \cup (R \circ R) \cup (R \circ R \circ R) \cup \cdots$, where $\circ$ is the composition of binary relations). By the reflexivity of R and the definition of its symmetric part, the structure $(X, \mathrm{sp}(R))$ is a *graph* (with loops), and the *cliques* in it are its maximal complete subgraphs. By the reflexivity of R and the definition of its transitive closure, the structure $(X, \mathrm{tc}(R))$ is a *quasi-ordered set*, and an *order filter* in it is any subset $A \subseteq X$ such that, for all $x, y \in X$, if $(x, y) \in \mathrm{tc}(R)$ and $x \in A$, then $y \in A$ too.

PROPOSITION 4. *Let X; $T \subseteq X^X$; R, $\mathrm{sp}(R)$, $\mathrm{tc}(R) \subseteq X^2$; and ${}^{\triangleright}\mathcal{A}$, ${}^{\triangleleft}\mathcal{A} \subseteq 2^X$ be as defined above.*
(i) Family ${}^{\triangleright}\mathcal{A}$ is an order ideal in the partially ordered set $(2^X, \subseteq)$. Its maximal members are the cliques of the graph $(X, \mathrm{sp}(R))$.
(ii) Family ${}^{\triangleleft}\mathcal{A}$ is a sub-lattice of lattice $(2^X, \cap, \cup)$. Its members are the order filters in the quasi-ordered set $(X, \mathrm{tc}(R))$. [5]

Proof. (i) That ${}^{\triangleright}\mathcal{A}$ is an order ideal by set inclusion (i.e., $A \in {}^{\triangleright}\mathcal{A}$ and $B \subseteq A$ implies $B \in {}^{\triangleright}\mathcal{A}$) is tantamount to Proposition 2.i. That the maximal members of ${}^{\triangleright}\mathcal{A}$ are the cliques of $(X, \mathrm{sp}(R))$ follows from Proposition 1.i (${}^{\triangleright}\mathcal{A}$ is the set of complete subgraphs of the graph). (ii) The first statement in this part is tantamount to Proposition 2.iii. To prove the second statement, first note that $R(A) \subseteq A$ if and only if $\mathrm{tc}(R)(A) \subseteq A$, for all $A \subseteq X$, as the "if" is simply due to inclusion $R \subseteq \mathrm{tc}(R)$, and the "only if" follows from Proposition 2.v. Thus ${}^{\triangleleft}\mathcal{A} = \{A \in 2^X : \mathrm{tc}(R)(A) \subseteq A\}$, and this is precisely the set of order filters in the quasi-ordered set $(X, \mathrm{tc}(R))$. □

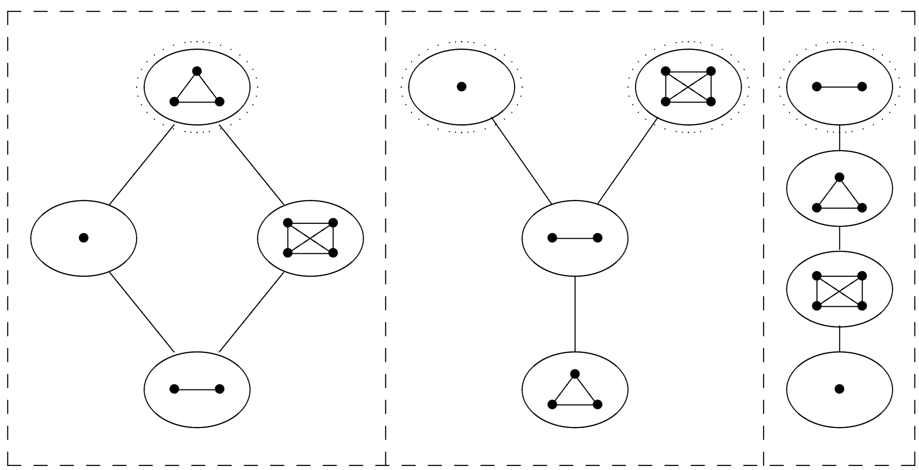

FIGURE 3. Three families of subsets (enclosed by solid, dotted, and dashed line contours) of a quasi-ordered set of 30 elements (the dots). Lines between the dots represent the symmetric part of the quasi-order (an equivalence).

The general properties in Proposition 4 undergo specialization when additional conditions are given concerning the transformation set T (and the adjacency R it determines). (i) If R is a *quasi-order* (which is the case, for example, when T is a *semigroup* and contains the identity transformation), then $\mathrm{sp}(R) = R \cap R^{-1}$ is an equivalence, and $\mathrm{tc}(R) = R$. In this situation three

[5]The link between a directed adjacency $R \subseteq X^2$ and the corresponding family of sets ${}^{\triangleleft}\mathcal{A} \subseteq 2^X$ is quite similar to that between a "surmise relation" and the corresponding "knowledge structure" of a "quasi-ordinal space", as defined in knowledge space theory [Doignon, Falmagne, 1999, p. 39]. Proposition 4.ii corresponds to Birkhoff's theorem [Birkhoff, 1937] which is called for in knowledge space theory.

special families of subsets of X may be distinguished: one is $X/\mathrm{sp}(R)$, i.e., the quotient set of X modulo equivalence $\mathrm{sp}(R)$; the second is the collection of blocks in $X/\mathrm{sp}(R)$ that are maximal in the quasi-order R; and the third is $X/\mathrm{tc}(\mathrm{sc}(R))$, i.e., the quotient set of X modulo the transitive closure of $\mathrm{sc}(R) = R \cup R^{-1}$ ($\mathrm{sc}(R)$ is the symmetric completion of R). This distinction is illustrated in Figure 3, in which the subsets forming the first, second and third family are those defined by contours of solid, dotted and dashed lines, respectively. Under the condition stated above (i.e., when R is a quasi-order) the sets in the first family are the maximal members of the order ideal ${}^{\rhd}\mathcal{A}$, and those in the second are the atoms in the lattice ${}^{\lhd}\mathcal{A}$. Owing to Proposition 2.vi, these latter sets are the members of the intersected family ${}^{\rhd}\mathcal{A} \cap {}^{\lhd}\mathcal{A}$, i.e., they are the subsets of X having T as both an upper and a lower bound of diversity. (ii) If T is a set (not necessarily a group) of *permutations*, then the adjacency R is *part-wise symmetric* (this may be proved "by contradiction" and taking account of the assumption that X has finite cardinality). This property combined with Proposition 2.iv implies that the family ${}^{\lhd}\mathcal{A}$ is closed under complementation. In turn, this property combined with Proposition 4.ii implies that ${}^{\lhd}\mathcal{A}$ is a field of subsets of X [6]. (iii) If T is a *permutation group*, then R is an *equivalence* and the whole system neatly simplifies, as $\mathrm{sp}(R) = \mathrm{tc}(R) = \mathrm{sc}(R) = R$. In this case, the maximal members of the order ideal ${}^{\rhd}\mathcal{A}$ are the blocks in the quotient set X/R, which are also the atoms in the field of sets ${}^{\lhd}\mathcal{A}$, so that ${}^{\rhd}\mathcal{A} \cap {}^{\lhd}\mathcal{A} = X/R$. Thus, ${}^{\lhd}\mathcal{A}$ is the family of "sets invariant under T" and ${}^{\rhd}\mathcal{A} \cap {}^{\lhd}\mathcal{A}$ in the family of "T-orbits", in the standard senses mentioned in the Introduction.

3. A SET OF TRANSFORMATIONS AND A PARTITION OF THEIR DOMAIN

The concepts $\rhd$ and $\lhd$ may be consistently extended from subsets of X to families of subsets of X. We are especially interested in families that are *partitions* of X, and we denote by $\mathbb{P}(X)$ their collection.

DEFINITION 2. *Let X be a set, $T \subseteq X^X$, and $\mathcal{P} \in \mathbb{P}(X)$. Set T represents an* upper bound *(symbol $T \rhd \mathcal{P}$) or* lower bound *(symbol $T \lhd \mathcal{P}$) of the diversity in the parts of the partition $\mathcal{P}$ depending on whether the first or the second of these two conditions holds true:*

$$T \rhd P \text{ for all } P \in \mathcal{P}$$
$$T \lhd P \text{ for all } P \in \mathcal{P}.$$

Thus, $T \rhd \mathcal{P}$ or $T \lhd \mathcal{P}$ depending on whether $\mathcal{P} \subseteq {}^{\rhd}\mathcal{A}$ or $\mathcal{P} \subseteq {}^{\lhd}\mathcal{A}$. For an example, let us refer to the transformation set T described by (1), and consider $\mathcal{P}_1 = \{\{b\}, \{b, c, d\}, \{e, f\}, \{g, h, i\}, \{j\}, \{k, l\}\}$, $\mathcal{P}_2 = \{\{a, b, c, d, e, f\}, \{g, h, i, j, k, l\}\}$, and $\mathcal{P}_3 = \{\{a, b, c, d\}, \{e, f\}, \{g, h, i, j\}, \{k, l\}\}$ (see Figure 1). Applying Definition 2 we see that $T \rhd \mathcal{P}_1$ but not $T \lhd \mathcal{P}_1$; $T \lhd \mathcal{P}_2$ but not $T \rhd \mathcal{P}_2$; and neither $T \rhd \mathcal{P}_3$ nor $T \lhd \mathcal{P}_3$. There is no partition $\mathcal{P}$ in this example such that both $T \rhd \mathcal{P}$ and $T \lhd \mathcal{P}$ [7].

A property of partitions is their one-to-one correspondence with equivalence relations. Specifically, if $\mathcal{P} = \{P_1, \ldots, P_n\}$ is a partition of X, then $\mathcal{P} = X/E$ with E the equivalence thus defined:

$$E = P_1^2 \cup \cdots \cup P_n^2.$$

This correspondence allows for simple characterizations of the concepts in Definition 2.

PROPOSITION 5. *Let $T \subseteq X^X$, $\mathcal{P} \in \mathbb{P}(X)$, $R \subseteq X^2$ be the adjacency determined by T, and $E \subseteq X^2$ be the equivalence corresponding to $\mathcal{P}$. Then:*
(i) $T \rhd \mathcal{P}$ if and only if $R \supseteq E$;
(ii) $T \lhd \mathcal{P}$ if and only if $R \subseteq E$.

[6] "Field of subsets of X" is a name used for a Boolean algebra of subsets of X, i.e., a set of subsets of X that is closed under union and complementation [Birkhoff, 1967, p. 12].

[7] The relations $\rhd$ and $\lhd$ may further be extended from partitions to functions. Specifically, if $T \subseteq X^X$, $f \in U^X$, and E_f is the equivalence determined by f on X (i.e., $(x, y) \in E_f$ iff $f(x) = f(y)$), then define $T \rhd f$ iff $T \rhd X/E_f$, and $T \lhd f$ iff $T \lhd X/E_f$. The properties of $\rhd$ and $\lhd$ for partitions may then be interpreted in relation to functions. For example, part (vii) of Proposition 6 means that if $f \in U^X$ and $g \in V^X$ are such that $T \rhd f$ and $T \lhd g$, then $X/E_f \preceq X/E_g$, which implies that there is an $h \in V^U$ such that $g = h \circ f$ (i.e., $T \rhd f$ and $T \lhd g$ for some $T \subseteq X^X$ implies that g is functionally dependent on f).

Proof. (i) Suppose $\mathcal{P} = \{P_1, \ldots, P_n\}$. Then $T \rhd \mathcal{P}$ iff ($T \rhd P_1$ and ... and $T \rhd P_n$) iff ($R \supseteq P_1^2$ and ... and $R \supseteq P_n^2$) iff $R \supseteq E$. (ii) Similarly, $T \lhd \mathcal{P}$ iff ($T \lhd P_1$ and ... and $T \lhd P_n$) iff ($R(P_1) \subseteq P_1$ and ... and $R(P_n) \subseteq P_n$) iff ($P_1 \times R(P_1) \subseteq P_1^2$ and ... and $P_n \times R(P_n) \subseteq P_n^2$) iff $R \subseteq E$ (as $\{P_1, \ldots, P_n\}$ is a partition of X, so $\{P_1 \times R(P_1), \ldots, P_n \times R(P_n)\}$ is a partition of R, and this justifies the last "iff"). □

Other elementary properties of partitions are the following: on the set $\mathbb{P}(X)$ of all partitions of X a binary relation $\preceq$ is defined such that $\mathcal{P} \preceq \mathcal{Q}$ (for any $\mathcal{P}, \mathcal{Q} \in \mathbb{P}(X)$) means that, for all $P \in \mathcal{P}$, there is a $Q \in \mathcal{Q}$ such that $P \subseteq Q$. Relation $\preceq$ is a partial order; more precisely, the partially ordered set $(\mathbb{P}(X), \preceq)$ is a lattice, which means that it is endowed with a meet operation $\wedge$ and a join operation $\vee$; for all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}(X)$, each part in the meet partition $\mathcal{P} \wedge \mathcal{Q}$ is the intersection of a part in $\mathcal{P}$ and a part in $\mathcal{Q}$ that are non-disjoint, and each part in the join partition $\mathcal{P} \vee \mathcal{Q}$ is the union of the parts in a connected component of the intersection graph of $\mathcal{P} \cup \mathcal{Q}$. These concepts are used in stating and proving the following properties of the relations $\rhd$ and $\lhd$ as referred to partitions.

PROPOSITION 6. *Let $T \subseteq X^X$, $R \subseteq X^2$, $\mathcal{P}, \mathcal{Q} \in \mathbb{P}(X)$, and $A \subseteq X$ be as supposed so far. Then the following implications hold true:*
(i) if $T \rhd \mathcal{P}$ and $\mathcal{Q} \preceq \mathcal{P}$, then $T \rhd \mathcal{Q}$;
(ii) if R is transitive, $T \rhd \mathcal{P}$, and $T \rhd \mathcal{Q}$, then $T \rhd \mathcal{P} \vee \mathcal{Q}$;
(iii) if $T \lhd \mathcal{P}$ and $\mathcal{P} \preceq \mathcal{Q}$, then $T \lhd \mathcal{Q}$;
(iv) if $T \lhd \mathcal{P}$ and $T \lhd \mathcal{Q}$, then $T \lhd \mathcal{P} \wedge \mathcal{Q}$;
(v) if $T \rhd \mathcal{P}$ and $T \lhd A$, then there are $P_1, \ldots, P_k \in \mathcal{P}$ such that $A = P_1 \cup \cdots \cup P_k$;
(vi) if $T \lhd \mathcal{P}$ and $T \rhd A$, then there is a $P \in \mathcal{P}$ such that $A \subseteq P$;
(vii) if $T \rhd \mathcal{P}$ and $T \lhd \mathcal{Q}$, then $\mathcal{P} \preceq \mathcal{Q}$.

Proof. (i) If $E_{\mathcal{P}}$ and $E_{\mathcal{Q}}$ are the equivalences corresponding to $\mathcal{P}$ and $\mathcal{Q}$, then the hypotheses imply $R \supseteq E_{\mathcal{P}} \supseteq E_{\mathcal{Q}}$, so that $T \rhd \mathcal{Q}$ (by Proposition 5.i). (ii) Each part B in the join partition $\mathcal{P} \vee \mathcal{Q}$ may be expressed as $B = A_1 \cup A_2 \cup \cdots \cup A_k$ with $\{A_1, \ldots, A_k\} \subseteq \mathcal{P} \cup \mathcal{Q}$ and $(A_1 \cup \cdots \cup A_{h-1}) \cap A_h \neq \emptyset$ for all $h = 2, \ldots, k$. The hypotheses and Proposition 2.ii (repeatedly applied) imply $T \rhd B$. This is true of all $B \in \mathcal{P} \vee \mathcal{Q}$, so $T \rhd \mathcal{P} \vee \mathcal{Q}$. (iii) The hypotheses imply $R \subseteq E_{\mathcal{P}} \subseteq E_{\mathcal{Q}}$, so that $T \lhd \mathcal{Q}$ (by Proposition 5.ii). (iv) The hypotheses imply $R \subseteq E_{\mathcal{P}}$ and $R \subseteq E_{\mathcal{Q}}$, so that $R \subseteq E_{\mathcal{P}} \cap E_{\mathcal{Q}} = E_{\mathcal{P} \wedge \mathcal{Q}}$, and $T \lhd \mathcal{P} \wedge \mathcal{Q}$ (by Proposition 5.ii). (v) The hypotheses imply $T \rhd P$ for all $P \in \mathcal{P}$ and $T \lhd A$, so that (either $P \subseteq A$ or $P \cap A = \emptyset$) for all $P \in \mathcal{P}$ (by Proposition 2.vi), and this amounts to the consequent of the implication. (vi) The hypotheses imply $T \lhd P$ for all $P \in \mathcal{P}$ and $T \rhd A$, so that (either $P \supseteq A$ or $P \cap A = \emptyset$) for all $P \in \mathcal{P}$ (by Proposition 2.vi), and this implies $A \subseteq P$ for one $P \in \mathcal{P}$ (the parts are disjoint). (vii) The hypotheses imply $R \supseteq E_{\mathcal{P}}$ and $R \subseteq E_{\mathcal{Q}}$, so that $E_{\mathcal{P}} \subseteq E_{\mathcal{Q}}$, which means $\mathcal{P} \preceq \mathcal{Q}$. □

For any fixed transformation set $T \subseteq X^X$, let ${}^{\rhd}\mathbb{P}$ and ${}^{\lhd}\mathbb{P}$ be the families of all partitions of X having T as an upper and, respectively, lower bound of diversity:

$$\begin{aligned} {}^{\rhd}\mathbb{P} &= \{\mathcal{P} \in \mathbb{P}(X) : T \rhd \mathcal{P}\} \\ {}^{\lhd}\mathbb{P} &= \{\mathcal{P} \in \mathbb{P}(X) : T \lhd \mathcal{P}\}. \end{aligned}$$

The next proposition specifies the algebraic profiles of these families. It does this by referring to the relations $\mathrm{sp}(R) = R \cap R^{-1}$ (i.e., the symmetric part of the adjacency R determined by T) and $\mathrm{tc}(\mathrm{sc}(R)) = \mathrm{tc}(R \cup R^{-1})$ (i.e., the transitive closure of the symmetric completion of R). The latter is an equivalence on X (it is the smallest equivalence including R), and when R is transitive then the former is itself an equivalence (it is the greatest equivalence included in R).

PROPOSITION 7. *Let $T \subseteq X^X$; R, $\mathrm{sp}(R)$, $\mathrm{tc}(\mathrm{sc}(R)) \subseteq X^2$; and ${}^{\rhd}\mathbb{P}, {}^{\lhd}\mathbb{P} \subseteq \mathbb{P}(X)$ be as defined above.*
(i) Family ${}^{\rhd}\mathbb{P}$ is an order ideal in the partially ordered set $(\mathbb{P}(X), \preceq)$. In particular, if R is transitive, then ${}^{\rhd}\mathbb{P}$ is an ideal in the lattice $(\mathbb{P}(X), \wedge, \vee)$, and the maximum in it is the quotient set $X/\mathrm{sp}(R)$.
(ii) Family ${}^{\lhd}\mathbb{P}$ is a filter in the lattice $(\mathbb{P}(X), \wedge, \vee)$. The minimum in it is the quotient set $X/\mathrm{tc}(\mathrm{sc}(R))$.

Proof. (i) That ${}^{\triangleright}\mathbb{P}$ is an order ideal follows from part (i) of Proposition 6, and that if R is transitive then ${}^{\triangleright}\mathbb{P}$ is an ideal in the partition lattice follows from parts (i) and (ii) of that proposition. Suppose that R is transitive (so that $\mathrm{sp}(R)$ is an equivalence) and put $\mathcal{Q} = X/\mathrm{sp}(R)$. Then $E_{\mathcal{Q}} = \mathrm{sp}(R) \subseteq R$, so that $\mathcal{Q} \in {}^{\triangleright}\mathbb{P}$ by Proposition 5.i. If $\mathcal{P}$ is any member of ${}^{\triangleright}\mathbb{P}$, then $P^2 \subseteq R$ for all $P \in \mathcal{P}$, so that for each $P \in \mathcal{P}$ there exists $Q \in \mathcal{Q}$ such that $P \subseteq Q$ (because each $Q \in \mathcal{Q}$ is a *maximal* subset of X such that $Q^2 \subseteq R$), hence $\mathcal{P} \preceq \mathcal{Q}$. Thus, in the presumed conditions, partition $\mathcal{Q}$ is the maximum member of the ideal ${}^{\triangleright}\mathbb{P}$. (ii) That ${}^{\triangleleft}\mathbb{P}$ is a filter in the partition lattice follows from parts (iii) and (iv) of Proposition 6. Put $\mathcal{Q} = X/\mathrm{tc}(\mathrm{sc}(R))$. For all $Q \neq Q' \in \mathcal{Q}$, $(Q \times Q') \cap R \subseteq (Q \times Q') \cap \mathrm{tc}(\mathrm{sc}(R)) = \emptyset$, so that $R(Q) \subseteq Q$, which means $\mathcal{Q} \in {}^{\triangleleft}\mathbb{P}$ by Proposition 1.ii and Definition 2. Consider any $\mathcal{P} \in {}^{\triangleleft}\mathbb{P}$ and suppose (to reach a contradiction) not $\mathcal{Q} \preceq \mathcal{P}$, which means that there are $Q \in \mathcal{Q}$ and $k \geq 2$ distinct blocks $P_1, \ldots, P_k$ in $\mathcal{P}$ such that $Q \subseteq P_1 \cup \cdots \cup P_k$ and $Q \cap P_h \neq \emptyset$ for all $h = 1, \ldots, k$. Then take any $x \in Q \cap P_1$ and $z \in Q \cap P_k$. Owing to the definition of $\mathcal{Q}$ and because $x, z \in Q$, there is a sequence $(y_1, \ldots, y_n)$ of elements of Q such that $y_1 = x$, $y_n = z$, and $(y_j, y_{j+1}) \in R$ or $(y_{j+1}, y_j) \in R$ for all $j = 1, \ldots, n-1$. Let i be any index such that $y_i \in P_1$ and $y_{i+1} \in P_h$ with $h \neq 1$ (such an index does exist, as $y_1 \in P_1$ and $y_n \in P_k$). Then $(y_i, y_{i+1}) \in R$ (which contradicts $T \triangleleft P_1$ implied by $\mathcal{P} \in {}^{\triangleleft}\mathbb{P}$) or $(y_{i+1}, y_i) \in R$ (which contradicts $T \triangleleft P_h$ also implied by $\mathcal{P} \in {}^{\triangleleft}\mathbb{P}$). These contradictory results imply that, if $\mathcal{P} \in {}^{\triangleleft}\mathbb{P}$ then $\mathcal{Q} \preceq \mathcal{P}$, so that the partition $\mathcal{Q}$ is the minimum member of the filter ${}^{\triangleleft}\mathbb{P}$. □

We complete Proposition 7 with comments on two special cases. (i) If R is a *transitive* relation (which is the case, for example, when T is a semigroup), then (X, R) is a quasi-ordered set and, by referring to Figure 3 as a generic example, the partitions $X/\mathrm{sp}(R)$ and $X/\mathrm{tc}(\mathrm{sc}(R))$ forming, respectively, the maximum of the ideal ${}^{\triangleright}\mathbb{P}$ and the minimum of the filter ${}^{\triangleleft}\mathbb{P}$ are illustrated by the subsets enclosed by, respectively, the solid line and the dashed line contours. Of course $\mathrm{sp}(R) \subseteq \mathrm{tc}(\mathrm{sc}(R))$, so that the maximum partition in ${}^{\triangleright}\mathbb{P}$ stands in the relation $\preceq$ with the minimum partition in ${}^{\triangleleft}\mathbb{P}$ (and note that, no matter whether R is transitive or not, $\mathcal{P} \preceq \mathcal{Q}$ for all $\mathcal{P} \in {}^{\triangleright}\mathbb{P}$ and $\mathcal{Q} \in {}^{\triangleleft}\mathbb{P}$, because of Proposition 6.vii). (ii) If R is a *transitive* and *symmetric* relation (which is the case, for example, when T is a permutation group), then $\mathrm{sp}(R) = R = \mathrm{tc}(\mathrm{sc}(R))$, so that the *same* partition X/R constitutes the maximum of the ideal ${}^{\triangleright}\mathbb{P}$ and the minimum of the filter ${}^{\triangleleft}\mathbb{P}$. If T is a permutation group, then X/R is the partition of X into T-orbits.

4. A HIERARCHY OF SETS OF TRANSFORMATIONS

So far we have referred to one set T of inner transformations, but the method we are presenting involves a sequence of such sets as a standard for rating diversity.

DEFINITION 3. *A* transformation hierarchy *on a basic set X is a sequence $\mathcal{T} = (T_1, \ldots, T_m)$ of transformation sets on X such that the sequence $(R_1, \ldots, R_m)$ of the adjacencies they determine satisfies these conditions:*

$R_1 = \mathrm{id}_X$ *(the identity relation on X)*

$R_m = X^2$ *(the universal relation on X)*

$R_i \subset R_{i+1}$ *for each $i = 1, \ldots, m-1$ (the sequence is strictly increasing).*

DEFINITION 4. *Let $\mathcal{T} = (T_1, \ldots, T_m)$ be a transformation hierarchy on X, and consider any $A \subseteq X$. The* outer *and, respectively,* inner diversity rank *of A relative to $\mathcal{T}$ (denoted by $O(A)$ and $I(A)$) are, respectively, the first number i in the sequence $(1, 2, \ldots, m)$ such that $T_i \triangleright A$ and the last number i in that sequence such that $T_i \triangleleft A$. In other words:*

$$O(A) = \min\{i = 1, \ldots, m : T_i \triangleright A\}$$
$$I(A) = \max\{i = 1, \ldots, m : T_i \triangleleft A\}.$$

When applying these definitions we use a hierarchy of sets of transformations (and invariance under those transformations) as a criterion for rating purposes. In these generic respects there is a resemblance between our method and theories using transformation sets for classification purposes (for example, for classifying geometric or perceptual properties according to permutation groups which leave them invariant, or for classifying measurement scales according to the automorphism groups of their numerical systems; [Klein, 1872/1893; Cassirer, 1944; Stevens, 1951; Luce, Krantz, Suppes, Tversky, 1990, ch. 22; van Gool, Moons, Pauwels, Wagemans, 1994]). There are, however, peculiarities of our method, in that the reference hierarchy may be formed of *arbitrary* sets of inner transformations (not necessarily permutation groups), and the specific target is to rate the overall *diversity* within sets of objects.

Conditions $R_1 = \mathrm{id}_X$ and $R_m = X^2$ in Definition 3 ensure that $T_1 \lhd A$ and $T_m \rhd A$, so that the ranks $I(A)$ and $O(A)$ do exist for all $A \subseteq X$. The members $T_{O(A)}$ and $T_{I(A)}$ of the reference hierarchy constitute, respectively, the least upper bound and the greatest lower bound of the diversity in set A. Parts (i) and (ii) of Proposition 3 imply that $T_i \rhd A$ for all $i \geq O(A)$, and $T_i \lhd A$ for all $i \leq I(A)$. Both schemes $O(A) \leq I(A)$ and $O(A) > I(A)$ are possible. The former implies $T_i \rhd A$ and $T_i \lhd A$ for each $i \in [O(A), I(A)]$ (the interval of integer numbers from $O(A)$ to $I(A)$), so that each transformation set T_i in the interval *precisely* characterizes the diversity in set A. The latter implies that neither $T_i \rhd A$ nor $T_i \lhd A$ for each $i \in [I(A)+1, O(A)-1]$.

We illustrate the above definitions by an example constructed on a set of simple pictures as shown in Figure 4. Each of the pictures represents a tree with a bush nearby, and they differ from one another in three variables: the form of the tree, the size of the tree, and the location of the bush. These variables have, respectively, four, three, and three possible values, denoted by integers from 1 to 4 as specified in the figure. Thus, the basic set X in this example is a set of $4 \times 3 \times 3 = 36$ pictures, determined by freely combining the values of the three variables. Each element x in the set may be coded as a triple of numbers (x_f, x_s, x_b), which are the values taken in it by the variables form, size, and bush. The triple recorded in each cell of Figure 4 is the code for the picture shown in it.

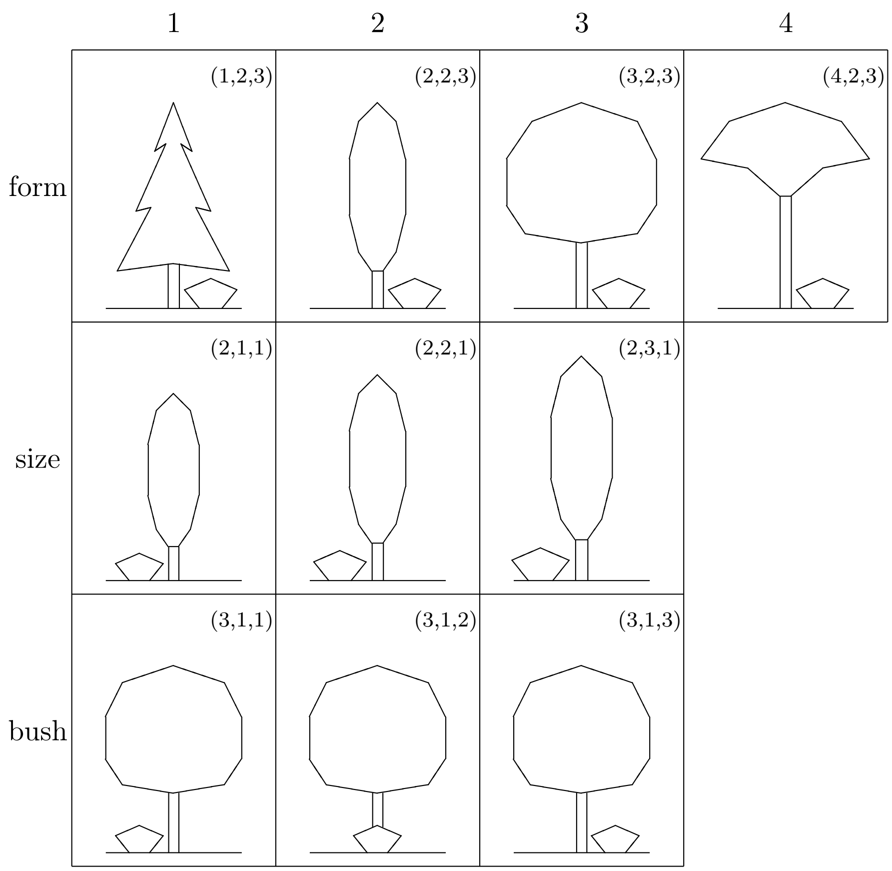

FIGURE 4. A sample of pictures illustrating the meaning of the possible values of the variables form, size, and bush.

In addition, let us suppose that for each variable some permutations on its range are determined *a priori*, to be interpreted as admissible substitution rules between the values of the variable. Specifically, let us presume the following sets of permutations:

$$\begin{aligned} F &= \{f_1 = 1234, f_2 = 2143, f_3 = 4321, f_4 = 3412\} \\ S &= \{s_1 = 123, s_2 = 213, s_3 = 132, s_4 = 321\} \\ B &= \{b_1 = 123, b_2 = 321, b_3 = 231, b_4 = 312\}. \end{aligned}$$

For example, $f_2 = 2143$, $s_4 = 321$, and $b_3 = 231$ mean the substitution rules $(1 \to 2,\ 2 \to 1,\ 3 \to 4,\ 4 \to 3)$, $(1 \to 3,\ 2 \to 2,\ 3 \to 1)$, and $(1 \to 2,\ 2 \to 3,\ 3 \to 1)$ acting within the ranges of the variables form, size, and bush, respectively. In these conditions, a transformation t_{ijk} acting within the basic set X of 36 pictures may be defined by combining the transformations f_i from F, s_j from S, and b_k from B. For example, t_{243} is the global transformation on X combining the local transformations f_2, s_4, and b_3, so that $t_{243}((1,1,1)) = (f_2(1), s_4(1), b_3(1)) = (2,3,2)$, $t_{243}((1,1,2)) = (f_2(1), s_4(1), b_3(2)) = (2,3,3)$, and so on.

The reference transformation sets (on X) we presume in this example are the following:

$$\begin{aligned} T_1 &= \{f_1\} \times \{s_1\} \times \{b_1\} \\ T_2 &= \{f_1\} \times \{s_1\} \times \{b_1, b_2\} \\ T_3 &= \{f_1\} \times \{s_1, s_2\} \times \{b_1, b_2\} \\ T_4 &= \{f_1, f_2\} \times \{s_1, s_2\} \times \{b_1, b_2\} \\ T_5 &= \{f_1, f_2\} \times \{s_1, s_2\} \times \{b_1, b_2, b_3, b_4\} \\ T_6 &= \{f_1, f_2\} \times \{s_1, s_2, s_3\} \times \{b_1, b_2, b_3, b_4\} \\ T_7 &= \{f_1, f_2, f_3\} \times \{s_1, s_2, s_3\} \times \{b_1, b_2, b_3, b_4\} \\ T_8 &= \{f_1, f_2, f_3\} \times \{s_1, s_2, s_3, s_4\} \times \{b_1, b_2, b_3, b_4\} \\ T_9 &= \{f_1, f_2, f_3, f_4\} \times \{s_1, s_2, s_3, s_4\} \times \{b_1, b_2, b_3, b_4\}. \end{aligned} \tag{2}$$

For example, T_4 is a set of $2 \times 2 \times 2 = 8$ inner transformations on X, i.e., all transformations which may be constructed by combining one of $\{f_1, f_2\}$ with one of $\{s_1, s_2\}$ and one of $\{b_1, b_2\}$ as stated above. The system $\mathcal{T} = (T_1, \ldots, T_9)$ thus defined is an increasing monotone sequence of sets of inner transformations of the basic set X, and each set T_h in the sequence collectively represents a similarity relation within the basic set X, a relation which is more inclusive or "permissive" the more advanced is the place of T_h in the sequence. An assumption we made in constructing the sequence is that a difference in the location of the bush (third variable) is less important than a difference in size (second variable), which in turn is less important than a difference in form (first variable). It is easily seen that sequence $\mathcal{T} = (T_1, \ldots, T_9)$ has the properties required by Definition 3, so that it may serve as a transformation hierarchy in rating diversity.

Having established this background, let us now consider any subset A of X, such as the sample of 8 pictures represented in Figure 5 and specified by these codes:

$$A = \{(1,1,1), (1,1,3), (1,2,1), (1,2,3), (2,1,1), (2,1,3), (2,2,1), (2,2,3)\}.$$

Through detailed examination it is seen that the transformation set T_4 is both the largest member of the hierarchy $\mathcal{T}$ such that $T_4 \lhd A$ (a step higher in the hierarchy is not possible, because, for example, t_{114} is in T_5, $(1,1,3)$ is in A, but $t_{114}((1,1,3)) = (1,1,2)$ is not in A), and the smallest member of the same hierarchy such that $T_4 \rhd A$ (a step lower is not possible, because, for example, $(1,1,1)$ and $(2,1,1)$ are both in A but there is no t_{ijk} in T_3 such that $(2,1,1) = t_{ijk}((1,1,1))$). Thus, according to Definition 4, number 4 is the inner $I(A)$ and outer $O(A)$ diversity rank of the object set A, as rated by referring to the hierarchy $\mathcal{T}$ of transformation sets. The following is another sample of 8 items from the same basic set:

$$B = \{(1,1,1), (1,1,3), (1,2,1), (1,2,3), (4,1,1), (4,1,3), (4,2,1), (4,2,3)\}.$$

Applying the same procedure, we can see that T_3 is the largest member of the hierarchy $\mathcal{T}$ such that $T_3 \lhd B$, and T_7 is the smallest member such that $T_7 \rhd B$, so that the inner rank $I(B) = 3$ and the outer rank $O(B) = 7$ are different in this case.

We resume the general discussion with a definition concerning partitions.

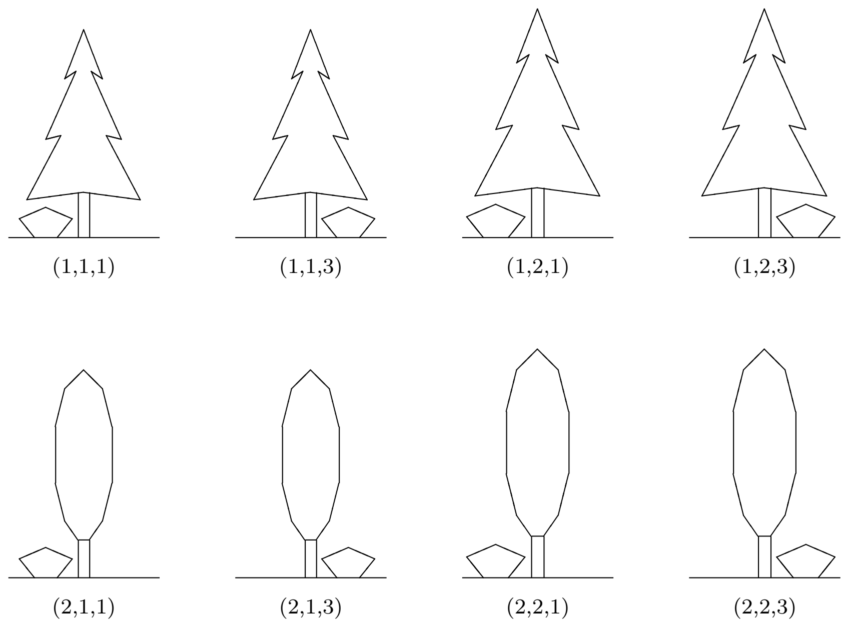

FIGURE 5. A sample of 8 pictures (trees-with-bush) to be rated in overall diversity.

DEFINITION 5. *Let* $\mathcal{T} = (T_1, \dots, T_m)$ *be a transformation hierarchy on* X, *and* $\mathcal{P}$ *a partition of* X. *The* outer *and, respectively,* inner diversity rank *of* $\mathcal{P}$ *relative to* $\mathcal{T}$ *(denoted by* $O(\mathcal{P})$ *and* $I(\mathcal{P})$*) are, respectively, the first number* i *in the sequence* $(1, 2, \dots, m)$ *such that* $T_i \rhd \mathcal{P}$ *and the last number* i *in that sequence such that* $T_i \lhd \mathcal{P}$.

Supposing that $\mathcal{P} = \{P_1, \dots, P_n\}$ and using Definition 2, the following equalities are obvious:

$$\begin{aligned} O(\mathcal{P}) &= \max(O(P_1), \dots, O(P_n)) \\ I(\mathcal{P}) &= \min(I(P_1), \dots, I(P_n)). \end{aligned}$$

Owing to the third condition in Definition 3, the scheme $O(\mathcal{P}) < I(\mathcal{P})$ is impossible for partitions (actually, if $O(\mathcal{P}) < I(\mathcal{P})$, then there would exist $O(\mathcal{P}) \le i < j \le I(\mathcal{P})$ such that $T_i \rhd \mathcal{P}$, $T_i \lhd \mathcal{P}$, $T_j \rhd \mathcal{P}$, and $T_j \lhd \mathcal{P}$, and because of Proposition 5 this would imply that $R_i = E_{\mathcal{P}} = R_j$, which contradicts the condition $R_i \subset R_j$). Thus $I(\mathcal{P}) \le O(\mathcal{P})$ for each partition $\mathcal{P}$ of X. The equality $I(\mathcal{P}) = O(\mathcal{P})$ is possible, and implies that $R_{I(\mathcal{P})}$ is an equivalence ($R_{I(\mathcal{P})} = E_{\mathcal{P}}$).

The next proposition collects a few simple properties of diversity ranks referring to subsets or partitions of X.

PROPOSITION 8. *Let* $\mathcal{T} = (T_1, \dots, T_m)$ *be a transformation hierarchy on* X, O *and* I *the diversity ranks (relative to* $\mathcal{T}$*),* A *any subset of* X, *and* $\mathcal{P}$ *and* $\mathcal{Q}$ *any partitions of* X. *Then these implications hold true:*
(i) if $\mathcal{P} \preceq \mathcal{Q}$, *then* $I(\mathcal{P}) \le I(\mathcal{Q})$ *and* $O(\mathcal{P}) \le O(\mathcal{Q})$;
(ii) if $O(\mathcal{P}) \le I(A)$, *then there are* $P_1, \dots, P_k \in \mathcal{P}$ *such that* $A = P_1 \cup \dots \cup P_k$;
(iii) if $O(A) \le I(\mathcal{P})$, *then there is a* $P \in \mathcal{P}$ *such that* $A \subseteq P$;
(iv) if $O(\mathcal{P}) \le I(\mathcal{Q})$, *then* $\mathcal{P} \preceq \mathcal{Q}$.

Proof. Part (i) is implied by parts (i) and (iii) of Proposition 6. If $O(\mathcal{P}) \le I(A)$, then there is a $T_i \in \mathcal{T}$ such that $T_i \rhd \mathcal{P}$ and $T_i \lhd A$, so that part (ii) is implied by part (v) of Proposition 6. Similarly, parts (iii) and (iv) are separately implied by parts (vi) and (vii) of that proposition. □

Implications (ii)-(iv) which have been proved in this way deserve notice as they show that knowledge of certain relations in diversity ranks between suitable items (information of ordinal type) enables us to infer set-theoretic relations between those items (information of structural type). Also note that, if $\mathcal{P}$ and $\mathcal{Q}$ are partitions of X such that $\mathcal{P} \preceq \mathcal{Q}$, then $I(\mathcal{P}) \leq I(\mathcal{Q})$ and $O(\mathcal{P}) \leq O(\mathcal{Q})$ owing to Proposition 8.i; furthermore, as $I(\mathcal{P}) \leq O(\mathcal{P})$ and $I(\mathcal{Q}) \leq O(\mathcal{Q})$ in general, then either one of these schemes must be true:

$$I(\mathcal{P}) \leq O(\mathcal{P}) \leq I(\mathcal{Q}) \leq O(\mathcal{Q}) \text{ or } I(\mathcal{P}) \leq I(\mathcal{Q}) < O(\mathcal{P}) \leq O(\mathcal{Q}).$$

We remark that not only the first, but also the second of them is compatible with $\mathcal{P} \preceq \mathcal{Q}$. To show this, let us suppose that $X = \{a, b, c, d, e\}$ is the basic set, T_i and T_{i+1} are two consecutive levels of a transformation hierarchy acting on it, the adjacencies R_i and R_{i+1} determined by them are equivalences, and $X/R_i = \{\{a, b\}, \{c\}, \{d\}, \{e\}\}$ and $X/R_{i+1} = \{\{a, b\}, \{c, d, e\}\}$ are the quotient sets modulo those equivalences. Further, consider the partitions $\mathcal{P} = \{\{a\}, \{b\}, \{c\}, \{d, e\}\}$ and $\mathcal{Q} = \{\{a, b, c\}, \{d, e\}\}$ of X. Then $\mathcal{P} \prec \mathcal{Q}$, $T_{i+1} \rhd \mathcal{P}$, not $T_i \rhd \mathcal{P}$, $T_i \lhd \mathcal{Q}$, and not $T_{i+1} \lhd \mathcal{Q}$ (by Proposition 5), so that $I(\mathcal{Q}) = i < i+1 = O(\mathcal{P})$ (by Definition 5). Scheme $I(\mathcal{P}) < I(\mathcal{Q}) < O(\mathcal{Q}) < O(\mathcal{P})$ is itself feasible, but implies that neither $\mathcal{P} \preceq \mathcal{Q}$ nor $\mathcal{Q} \preceq \mathcal{P}$ (owing to Proposition 8.i).

5. INDIRECT ACTION OF TRANSFORMATIONS

Addressing the problem of rating diversity in a transformational perspective makes it possible to use a transformation hierarchy $\mathcal{T} = (T_1, \ldots, T_m)$ for rating not only subsets of the domain X on which the transformations in $\mathcal{T}$ directly act, but also subsets of some other domain X^*, provided that there is a systematic relationship between X and X^*. This possibility rests on the process of "transformation induction", which means that in suitable conditions a transformation t acting within X uniquely determines a transformation t^* acting within X^* [8]. In the introduction we mentioned "actions of permutation groups" as a related standard concept of algebra. In footnote 3 we remarked that this possibility is the very reason why the adjacency R determined by a transformation set T cannot replace T everywhere in our analysis (in general, the transformation set T^* induced by T cannot be recovered from R).

In principle, there are several methods of transformation induction, as there are several kinds of relationship that can serve as a bridge for translating a transformation on a domain X into a transformation on another domain X^*. Here we discuss *induction by substitution* [9]. As a basic example, let us suppose that $X^* = X^n$ for a fixed positive integer n, that is, X^* is the set of all "words" of length n in the "alphabet" X. If t is any inner transformation on X, then an inner transformation t^* on X^* becomes determined as follows:

$$t^*(x^*) = (t(x_1), \ldots, t(x_n)), \text{ for all } x^* = (x_1, \ldots, x_n) \in X^*.$$

Thus, transformation t^* changes each n-tuple $x^* = (x_1, \ldots, x_n) \in X^*$ into an n-tuple $t^*(x^*) = (t(x_1), \ldots, t(x_n)) \in X^*$ by *substituting* each component x_j of x^* with the component $t(x_j)$ specified by t. More generally, if T is a set of inner transformations on X ($T \subseteq X^X$), then $T^* = \{t^* : t \in T\}$ is a set of inner transformations on X^* ($T^* \subseteq X^{*X^*}$), and the latter preserves some properties of the former. For example, if T is a semigroup, then T^* is also a semigroup, and is homomorphic to T.

In these conditions, for any subset A^* of X^* it appears proper to define the relations $\rhd$ and $\lhd$

[8]We are using "induction" in a sense similar to that used by Tarski, Givant [1987, p. 57] in discussing "logical objects".

[9]A different kind of transformation induction (by *composition*) may be considered when a function f from X to X^* is available. Then for any $t \in X^X$ the following set of pairs can be determined:

$$t^* = \{(x^*, y^*) \in X^{*2} : x^* = f(x) \text{ and } y^* = f(t(x)) \text{ for some } x \in X\}.$$

If f and t are such that $f(x) = f(y)$ implies $f(t(x)) = f(t(y))$ for all $x, y \in X$, then t^* is a function (i.e., $t^* \in X^{*X^*}$). In particular, if f is a bijective function, then t^* has the form $f \circ t \circ f^{-1}$ of a "conjugate" of the transformation t.

as follows:

$$T \rhd A^* \text{ iff } T^* \rhd A^*$$
$$T \lhd A^* \text{ iff } T^* \lhd A^*.$$

In other words, $T \rhd A^*$ (T represents an *upper bound* of the diversity in A^*) means that for all $x^*, y^* \in A^*$ a transformation $t \in T$ exists which is able to change x^* into y^* (so that $y^* = t^*(x^*)$), whereas $T \lhd A^*$ (T represents a *lower bound* of the diversity in A^*) means that, for all $x^* \in A^*$ and $t \in T$, the result $t^*(x^*)$ is itself in A^*. Other concepts discussed in the previous sections can also be extended in this manner – in particular, the diversity ranks $O(A^*)$ and $I(A^*)$ of a subset of X^*, as compared with a transformation hierarchy $\mathcal{T}$ on X. Note that this way of adapting the concepts makes sense only if the following assumption is accepted: each inner transformation in T expresses a similarity not only when *individual* elements of X are compared, but also when *combinations* of such elements are compared – i.e., for all $x^* = (x_1, \dots, x_n)$, $y^* = (y_1, \dots, y_n)$ in X^*, the availability of a $t \in T$ such that ($y_1 = t(x_1)$ and ... and $y_n = t(x_n)$) signifies that y^* is somehow similar or substitutable to x^*.

We illustrate this part of the method by an example which continues that presented in Section 4. Suppose that X^* is the set of all fourfold rows of trees-with-bush of the kind shown in Figure 4 (in set-theoretic terms, X^* is the Cartesian power X^4 of the basic set X of 36 trees-with-bush). The following expression specifies a sample of 6 such rows of trees, which are depicted in Figure 6:

$$\begin{aligned} A^* = \{ & a_1^* = ((3,1,1),(4,1,2),(2,2,3),(4,1,2)), a_2^* = ((4,1,1),(3,1,2),(1,2,3),(3,1,2)), \\ & a_3^* = ((3,2,1),(4,2,2),(2,1,3),(4,2,2)), a_4^* = ((3,1,3),(4,1,2),(2,2,1),(4,1,2)), \\ & a_5^* = ((4,1,3),(3,1,2),(1,2,1),(3,1,2)), a_6^* = ((3,2,3),(4,2,2),(2,1,1),(4,2,2))\}. \end{aligned}$$

Through detailed examination it is seen that T_2 is the largest member of hierarchy (2) such that $T_2 \lhd A^*$ (for example, $t_{112} \in T_2$, $t_{122} \in T_3 \setminus T_2$, $t_{112}(a_1^*) = a_4^* \in A^*$, but $t_{122}(a_2^*) = ((4,2,3),(3,2,2),(1,1,1),(3,2,2)) \notin A^*$), and T_4 is the smallest member of the same hierarchy such that $T_4 \rhd A^*$ (for example, $a_2^* = t_{211}(a_1^*)$ with $t_{211} \in T_4$, but there is no $t_{ijk} \in T_3$ such that $a_2^* = t_{ijk}(a_1^*)$). Thus, $I(A^*) = 2$ and $O(A^*) = 4$ are the inner and outer ranks of diversity of the set $A^* = \{a_1^*, ..., a_6^*\}$ of composite objects, when it is rated by referring to hierarchy (2) of transformation sets on the set X of primitive objects.

6. CONCLUDING COMMENTS

A peculiar characteristic of the method we discussed is the kind of output it produces: some set A of objects is considered and (in suitable conditions) the method yields an ordinal evaluation of the *overall diversity* within A. We believe that there may be situations of psychological research in which this kind of output has scientific importance. For example, suppose that X is a set of figural patterns related to one another by transformations of various kinds and/or degrees (some patterns are derivable from one another by simple and small changes, others by substantial and large changes, and there are also intermediate levels between these extremes). Also suppose that in an experiment one such pattern x is fixed as the "standard stimulus" and several other patterns from X are compared with x in separate trials. The task of the participant in a trial is to judge whether the "comparison stimulus" y presented in the trial is or is not the same (with tolerable perturbation) as the standard stimulus x (a same/different judgment task). Experiments of this kind have actually been performed in studying "perceptual shape equivalence" (e.g. [Niall, Macnamara, 1990; Wagemans, 1993; Wagemans, van Gool, Lamote, Foster, 2000]). The net result of such an experiment (when run on one participant) would be a subset A of X, i.e., the set of comparison stimuli y that received the response "same" when compared with the standard stimulus x. In such a context it would be scientifically important to be able to evaluate the overall diversity within set A, by referring to a hierarchy $\mathcal{T} = (T_1, \dots, T_m)$ of sets of transformations acting within the basic set X. That evaluation would represent (with approximation) the boundary between the transformations that do not disrupt the figural identity of the standard pattern, and all other transformations in the system.

In the introduction and elsewhere in this paper we suggested that our method amounts to a variation on the classic theme of using permutation groups for classification purposes. Consistent

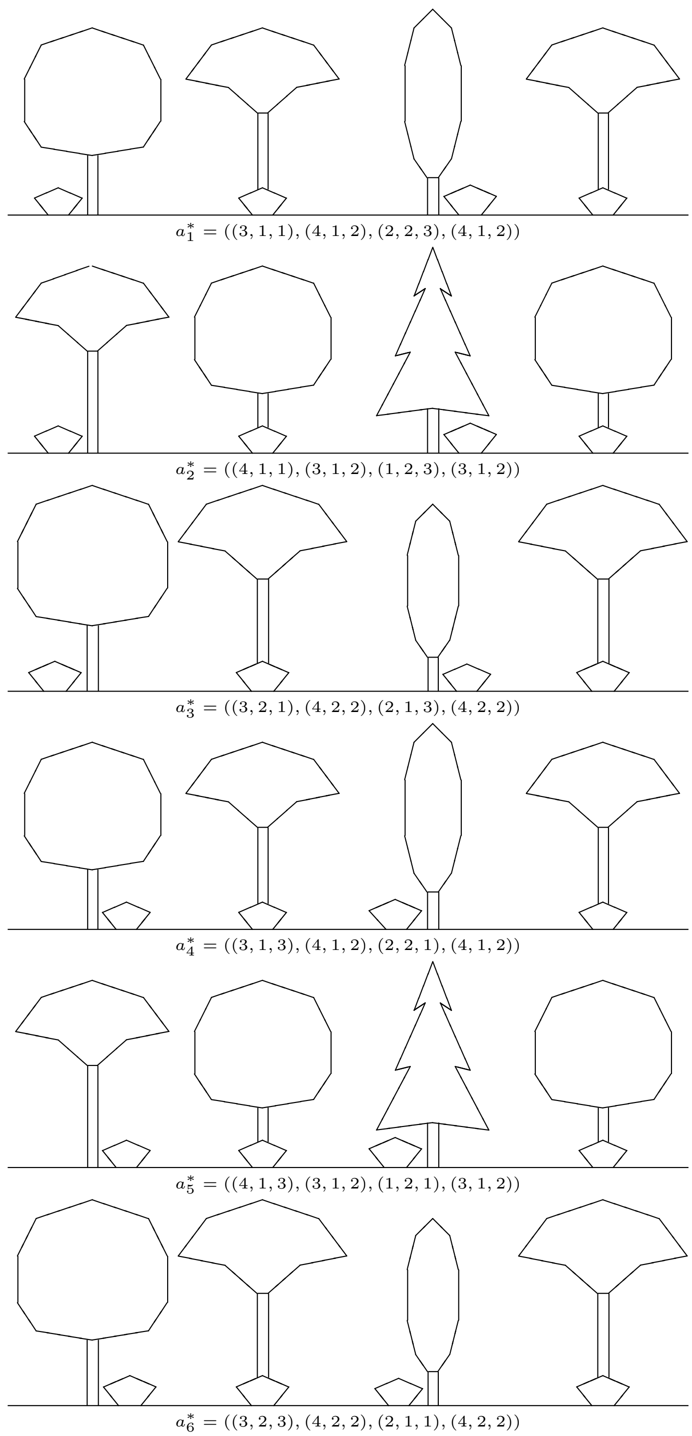

FIGURE 6. A sample of 6 fourfold rows of trees-with-bush.

with standard versions of this theme (e.g., in measurement theory) we referred to sets of inner transformations of a basic set X, interpreted them as substitution rules, and judged the diversity of any subset of X by comparing it with the diversity expressible by those transformations. The main difference in the premises of our approach is that the transformation sets may be any sets of inner transformations of X, and not necessarily permutation groups on X. This choice allows a freer context, and we stated a reason for it in the Introduction, when we noted the difference between similarity and equivalence relations. Of course, a greater freedom has a cost. In standard versions of the theme the transformation sets are permutation groups, and typically these are standard and well-known groups which constitute a kind of absolute reference system (e.g., in distinguishing the main types of measurement scales reference is made to the sets of bijective, increasing monotone, increasing linear, and dilation transformations on the real axis, and these are standard and well-known permutation groups on the axis). In contrast, when applying the suggested method to a specific empirical problem we may need to construct the intended transformation sets *ad hoc* for that problem (and, together with the object set A, these constitute the input of the method). When X is a small set, such a construction is feasible, as illustrated by the example in Section 4.

As there are several methods of evaluating similarities and dissimilarities in a psychological context, the following question is naturally raised: are there any special advantages of the transformational approach to the problem? This question concerns not only the modified version of the approach discussed in this paper, but also the standard version that only involves permutation groups. We answer the question by indicating two characteristics of the transformational approach. One is its *set-theoretic character*, which means that the essential basis of the procedure is a hierarchy of sets of inner transformations of a basic set, each transformation being interpreted as a substitution rule in that set. In a sense, the set-theoretic character is a guarantee of generality, because transformations as substitution rules may in principle be defined within any kind of basic set, irrespective of the quality or complexity of the items constituting it. For example, the method could be applied for rating the overall diversity within a set of shapes, or combinations of shapes, or pictures of faces, etc., once a suitable reference system of inner transformations has been defined on the relevant basic set. The other characteristic relates to the *indirect action* of inner transformations, which means that a transformation acting within a set may induce a transformation acting within another set (if the latter set is systematically related to the former), so that the original transformation also acts within this other set in a mediated way. In our analysis we illustrated this capacity on the set of partitions of a basic set (Section 3) and on a set of objects constructed by combining elements of a basic set (Section 5). The stated property enables the *same* system of inner transformations of a basic set to serve as a *common* criterion for rating the diversity not only within subsets of that basic set, but also within subsets of other domains connected with it. On this account that property may be counted as an advantage of the transformational approach[10].

REFERENCES

BARTHÉLEMY J.-P., GUÉNOCHE A. (1991), *Trees and proximity representations*, Chichester (UK), Wiley.

BIRKHOFF G. (1937), "Rings of sets", *Duke Mathematical Journal* 3, pp. 443-454.

BIRKHOFF G. (1967), *Lattice theory*, Providence (RI), American Mathematical Society.

CASSIRER E. (1944), "The concept of group and the theory of perception", *Philosophy and Phenomenological Research* 5, pp. 1-36.

DIXON J.D., MORTIMER B. (1996), *Permutation groups*, New York, Springer.

DOIGNON J.-P., FALMAGNE J.-C. (1999), *Knowledge spaces*, Berlin, Springer.

HAHN U., CHARTER N., RICHARDSON L.B. (2003), "Similarity as transformation", *Cognition* 87, pp. 1-32.

KLEIN F. (1872-1893), "Vergleichende Betrachtungen über neuere geometrische Forschungen ('Erlanger Programm')", *Mathematische Annalen* 43, pp. 63-100.

[10]The property in question is effectively illustrated in measurement theory, in which the group of automorphisms of the numerical relational system of a measurement scale serves as the common benchmark for judging the meaningfulness or appropriateness of several kinds of entities relating to the values of the scale, such as statements on the values, statistics computable on the values, statements on the values of the statistics, etc. [Pfanzagl, 1971; Luce *et al.*, 1990, ch. 22].

LUCE R.D., KRANTZ D.H., SUPPES P., TVERSKY A. (1990), *Foundations of measurement*, Vol. 3. *Representation, axiomatization, and invariance*, New York, Academic Press.

MUNDY J.L., ZISSERMAN A. (eds.) (1992), *Geometric invariance in computer vision*, Cambridge (MA), MIT Press.

NIALL K.K., MACNAMARA J. (1990), "Projective invariance and picture perception", *Perception* 19, pp. 637-660.

PFANZAGL J. (1971), *Theory of measurement*, Würzburg, Physica-Verlag.

STEVENS S.S. (1951), "Mathematics, measurement, and psychophysics", S.S. Stevens (ed.), *Handbook of experimental psychology*, New York, Wiley, pp. 1-49.

TARSKI A., GIVANT S. (1987), *A formalization of set theory without variables*, Providence (RI), American Mathematical Society.

TVERSKY A. (1977), "Features of similarity", *Psychological Review* 84, pp. 327-352.

VAN GOOL L., MOONS T., PAUWELS E., WAGEMANS J. (1994), "Invariance from the Euclidean geometer's perspective", *Perception* 23, pp. 547-561.

WAGEMANS J. (1993), "Skewed symmetry: A nonaccidental property used to perceive visual forms", *Journal of Experimental Psychology: Human Perception and Performance* 19, pp. 364-380.

WAGEMANS J., VAN GOOL L., LAMOTE C., FOSTER D.H. (2000), "Minimal information to determine affine shape equivalence", *Journal of Experimental Psychology: Human Perception and Performance* 26, pp. 443-468.